AN OPTICAL PHYSICS INSPIRED CNN APPROACH FOR INTRINSIC IMAGE DECOMPOSITION

Harshana Weligampola1 Gihan Jayatilaka2 Suren Sritharan1 Parakrama Ekanayake2 Roshan Ragel5 Vijitha Herath6 Roshan Godaliyadda2
1Faculty of IT and Computing, Sri Lanka Technological Campus, Padukka [10500], Sri Lanka
2Faculty of Engineering, University of Peradeniya [20400], Sri Lanka

Overview
Generating reflectance and shading from a single image is a challenging task when there is no ground truth. We propose a novel reflectance approximation map to train the neural network and a physics-based loss function to learn intrinsic properties in an image. Through numerical evaluation metrics, we show that the proposed model performs consistently well with different datasets consist of variety of scenes. There is room for improvement in regards to the color leakage problem in the shading map.

Problem
Intrinsic Image Decomposition (IID) is the problem of decomposing an image into its constituents

Image Interpretation

The perceived pixel intensity of an image is given by the Phong's model. Main components of an image:
Ambient Diffused Specular

Input image

Network architecture

Loss functions

Reconstruction loss

Shading smoothness loss

Reflectance Approximation Map (RAM) loss

results

Table: Evaluation Metrics

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSE</th>
<th>PSNR</th>
<th>SSIM</th>
<th>NIQE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letry et al.</td>
<td>21.87</td>
<td>35.28</td>
<td>0.96</td>
<td>7.55</td>
</tr>
<tr>
<td>CGIntrinsic</td>
<td>63.28</td>
<td>18.95</td>
<td>0.36</td>
<td>14.78</td>
</tr>
<tr>
<td>Retinex-net</td>
<td>4.88</td>
<td>34.64</td>
<td>0.90</td>
<td>7.63</td>
</tr>
<tr>
<td>Ours</td>
<td>2.00</td>
<td>43.12</td>
<td>0.95</td>
<td>7.63</td>
</tr>
</tbody>
</table>

Consider a narrow band and assuming that ambient illumination is constant, only one light source exists, and specular term is negligible,

\[ I_p(\lambda) = r_p(\lambda) \left[ \bar{I} \cdot \hat{N} \right] d(\lambda) \]

Similarly we can get the RAM for blue and green channels as well.

\[ M_{\text{RAM}} = \left[ m_{\text{R}}^q \right] = \frac{\left[ \bar{J}(\lambda_s, \lambda_p) + \bar{J}(\lambda_s, \lambda_g) \right]}{2} \quad \text{if} \quad c = R \]

\[ \text{Shading output} \quad S \]

\[ \text{Reflection output} \quad R \]

<table>
<thead>
<tr>
<th>Loss functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \mathcal{L} = \alpha_1 \mathcal{L}<em>{\text{recon}} + \alpha_2 \mathcal{L}</em>{\text{ss}} + \alpha_3 \mathcal{L}<em>{\text{reg}} + \alpha_4 \mathcal{L}</em>{\text{eg}} + \alpha_5 \mathcal{L}_{\text{ram}} ]</td>
</tr>
</tbody>
</table>

\[ \mathcal{L}_{\text{recon}} = \| R \mathbf{S} - \mathbf{I} \|_1 \]

\[ \mathcal{L}_{\text{ss}} = \| \nabla \log I_p(\lambda) \times \nabla \log I_p(\lambda) \|_1 \]

\[ \mathcal{L}_{\text{reg}} = \| f_{\text{reg}}(R \mathbf{S}) - f_{\text{reg}}(\mathbf{I}) \|_1 \]

\[ \mathcal{L}_{\text{eg}} = \| f_{\text{eg}}(R \mathbf{S}) - f_{\text{eg}}(\mathbf{I}) \|_1 \]

\[ \mathcal{L}_{\text{ram}} = \| (R \mathbf{S} - f_{\text{ram}}(\mathbf{I})) \times f_{\text{ram}}(\mathbf{I}) \|_1 \]